



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification 7 :  <b>C12N 15/31, C07K 14/40, G01N 33/50,  C12Q 1/68, A61K 31/7088, 38/16, C07K  16/14</b></p>	<p><b>A2</b></p>	<p>(11) International Publication Number: <b>WO 00/34481</b></p>
<p>(21) International Application Number: <b>PCT/EP99/09833</b></p>	<p>(22) International Filing Date: <b>6 December 1999 (06.12.99)</b></p>	<p>(43) International Publication Date: <b>15 June 2000 (15.06.00)</b></p>
<p>(30) Priority Data:  <b>98204122.0 4 December 1998 (04.12.98) EP</b></p>	<p>(71) Applicant (for all designated States except US): <b>JANSSEN PHARMACEUTICA N.V. [BE/BE]; Turnhoutseweg 30, B-2340 Beerse (BE).</b></p> <p>(72) Inventors; and  (75) Inventors/Applicants (for US only): <b>CONTRERAS, Roland, Henri [BE/BE]; University of Gent, K.L. Ledeganckstraat 35, B-9000 Gent (BE). NELISSEN, Bart [BE/BE]; Janssen Pharmaceutica N.V., Turnhoutseweg 30, B-2340 Beerse (BE). DE BACKER, Marianne, Denise [BE/BE]; Janssen Pharmaceutica N.V., Turnhoutseweg 30, B-2340 Beerse (BE). LUYTEN, Walter, Herman, Maria, Louis [BE/BE]; Janssen Pharmaceutica N.V., Turnhoutseweg 30, B-2340 beerse (BE). VIAENE, Jasmine, Elza [BE/BE]; University of Gent, K.L. Ledeganck Straat 35, B-9000 Gent (BE). LOGGHE, Marc, George [BE/BE]; University of Gent, K.L. Ledeganckstraat 35, B-9000 Gent (BE). VIALARD,</b></p>	<p>(74) Agent: <b>BOULT WADE TENNANT; 27 Furnival Street, London EC4A 1PQ (GB).</b></p>
<p>(71) Applicant (for all designated States except US): <b>JANSSEN PHARMACEUTICA N.V. [BE/BE]; Turnhoutseweg 30, B-2340 Beerse (BE).</b></p>		<p>(81) Designated States: <b>AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</b></p>
<p>(54) Title: <b>DRUG TARGETS IN CANDIDA ALBICANS</b></p>	<p>(57) Abstract</p>	<p><b>Published</b>  Without international search report and to be republished upon receipt of that report.</p>
<p>The present invention is concerned with the identification of genes or functional fragments thereof from <i>Candida albicans</i> which are critical for growth and cell division and which genes may be used as selective drug targets to treat <i>Candida albicans</i> associated infections. Novel nucleic acid sequences from <i>Candida albicans</i> are also provided and which encode the polypeptides which are critical for growth of <i>Candida albicans</i>. Methods for the identification of anti-fungal compounds which inhibit fungal or yeast growth are also contemplated.</p>		
<p>ATTORNEY DOCKET NUMBER: 10182-016-999  SERIAL NUMBER: 10/032,585  REFERENCE: <b>BX</b></p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

- 1 -

**DRUG TARGETS IN CANDIDA ALBICANS**

The present invention is concerned with the identification of genes or functional fragments thereof from *Candida albicans* which are critical for growth and cell division and which genes may be used as selective drug targets to treat *Candida albicans* associated infections. Novel nucleic acid sequences from *Candida albicans* are also provided and which encode the polypeptides which are critical for growth of *Candida albicans*.

Opportunistic infections in immunocompromised hosts represent an increasingly common cause of mortality and morbidity. *Candida* species are among the most commonly identified fungal pathogens associated with such opportunistic infections, with *Candida albicans* being the most common species. Such fungal infections are thus problematical in, for example, AIDS populations in addition to normal healthy women where *Candida albicans* yeasts represent the most common cause of vulvovaginitis.

Although compounds do exist for treating such disorders, such as, amphotericin, these drugs are generally limited in their treatment because of their toxicity and side effects. Therefore, there exists a need for new compounds which may be used to treat *Candida* associated infections in addition to compounds which are selective in their action against *Candida albicans*.

Classical approaches for identifying anti-fungal compounds have relied almost exclusively on inhibition of fungal or yeast growth as an endpoint. Libraries of natural products, semi-synthetic, or synthetic chemicals are screened for their ability to kill or arrest growth of the target pathogen or a related nonpathogenic model organism. These tests are

- 2 -

cumbersome and provide no information about a compound's mechanism of action. The promising lead compounds that emerge from such screens must then be tested for possible host-toxicity and detailed  
5 mechanism of action studies must subsequently be conducted to identify the affected molecular target.

The present inventors have now identified a range of nucleic acid sequences from *Candida albicans* which encode polypeptides which are critical for its  
10 survival and growth. These sequences represent novel targets which can be incorporated into an assay to selectively identify compounds capable of inhibiting expression of such polypeptides and their potential use in alleviating diseases or conditions associated  
15 with *Candida albicans* infection.

Therefore, according to a first aspect of the invention there is provided a nucleic acid molecule encoding a polypeptide which is critical for survival and growth of the yeast *Candida albicans* and which  
20 nucleic acid molecule comprises any of the sequences of nucleotides illustrated in any of Sequence ID Nos. 1 to 9.

Whilst the molecules defined herein have been established as being critical for growth and  
25 metabolism of *Candida albicans*, for some of the molecules no apparent functionality has been assigned by virtue of the fact that no functionally related sequences in other prokaryotic or eukaryotic organism can be found in respective databases. Thus,  
30 advantageously these sequences may be species specific in which case they may be used as selective targets for treatment of diseases mediated by *Candida Albicans* infection. Thus, in one aspect of the invention the nucleic acid molecules preferably  
35 comprise the sequences identified in sequence ID Nos. 1, 4, 5 to 9.

- 3 -

In another aspect of the invention the sequences have been arranged functionally and of nucleotides illustrated in Sequence ID Nos. 2 or 3 are preferred and even more preferably in Sequence ID No. 2 and fragments or derivatives of said nucleic acid molecules.

Letters utilised in the sequences according to the invention which are not recognisable as letters of the genetic code signify a position in the nucleic acid sequence where one or more of bases A, G, C or T can occupy the nucleotide position. Representative letters used to identify the range of bases which can be used are as follows:

15	M:	A or C
	R:	A or G
	W:	A or T
	S:	C or G
	Y:	C or T
20	K:	G or T
	V:	A or C or G
	H:	A or C or T
	D:	A or G or T
	B:	C or G or T
25	N:	G or A or T or C

In one embodiment of each of the above identified aspects of the invention the nucleic acid may comprise a mRNA molecule or alternatively a DNA and preferably a cDNA molecule.

Also provided by the present invention is a nucleic acid molecule capable of hybridising to the nucleic acid molecules illustrated in any of Figures 1 to 9 under high stringency conditions such as antisense molecule and which conditions are generally known to those of skill in the art.

Stringency of hybridisation as used herein refers

- 4 -

to conditions under which polynucleic acids are stable. The stability of hybrids is reflected in the melting temperature ( $T_m$ ) of the hybrids.  $T_m$  can be approximated by the formula:

5

$$81.5^{\circ}\text{C} + 16.6(\text{Log}_{10}[\text{Na}^+] + 0.41 (\% \text{G\&C}) - 600\text{L}/\text{L}$$

wherein L is the length of the hybrids in nucleotides.  $T_m$  decreases approximately by 1-1.5°C with every 1% decrease in sequence homology.

10

The term "stringency" refers to the hybridisation conditions wherein a single-stranded nucleic acid joins with a complementary strand when the purine or pyrimidine bases therein pair with their corresponding base by hydrogen bonding. High stringency conditions favour homologous base pairing whereas low stringency conditions disfavour non-homologous base pairing.

15

"Low stringency" conditions comprise, for example, a temperature of about 37°C or less, a formamide concentration of less than about 50%, and a moderate to low salt (SSC) concentration; or, alternatively, a temperature of about 50°C or less, and a moderate to high salt (SSPE) concentration, for example 1M NaCl.

20

"High stringency" conditions comprise, for example, a temperature of about 42°C or less, a formamide concentration of less than about 20%, and a low salt (SSC) concentration; or, alternatively, a temperature of about 65°C, or less, and a low salt (SSPE) concentration. For example, high stringency conditions comprise hybridization in 0.5 M  $\text{NaH}_2\text{PO}_4$ , 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C (Ausubel, F.M. et al. Current Protocols in Molecular Biology, Vol. I, 1989; Green Inc. New York, at 2.10.3).

30

35

"SSC" comprises a hybridization and wash solution. A stock 20X SSC solution contains 3M sodium

- 5 -

chloride, 0.3M sodium citrate, pH 7.0.

"SSPE" comprises a hybridization and wash solution. A 1X SSPE solution contains 180 mM NaCl, 9mM Na<sub>2</sub>HPO<sub>4</sub> and 1 mM EDTA, pH 7.4.

5       The nucleic acid capable of hybridising to nucleic acid molecules according to the invention will generally be at least 70%, preferably at least 80 or 90% and more preferably at least 95% homologous to the nucleotide sequences illustrated in any of Figures 1  
10      to 9.

      The DNA molecules according to the invention may, advantageously, be included in a suitable expression vector to express polypeptides encoded therefrom in a suitable host which are critical for growth and  
15      survival of *Candida albicans*.

      An expression vector according to the invention includes a vector having a nucleic acid according to the invention operably linked to regulatory sequences, such as promoter regions, that are capable of  
20      effecting expression of said DNA fragments. The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. Such vectors may be transformed into a suitable host  
25      cell to provide for expression of a polypeptide according to the invention. Thus, in a further aspect, the invention provides a process for preparing polypeptides according to the invention which comprises cultivating a host cell, transformed,  
30      transfected or infected with an expression vector as described above under conditions to provide for expression by the vector of a coding sequence encoding the polypeptides, and recovering the expressed polypeptides.

35      The vectors may be, for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression

- 6 -

of said nucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable markers, such as, for example, ampicillin resistance.

5 Regulatory elements required for expression include promoter sequences to bind RNA polymerase and transcription initiation sequences for ribosome binding. For example, a bacterial expression vector may include a promoter such as the lac promoter and  
10 for translation initiation the Shine-Dalgarno sequence and the start codon AUG. Similarly, a eukaryotic expression vector may include a heterologous or homologous promoter for RNA polymerase II, a downstream polyadenylation signal, the start codon  
15 AUG, and a termination codon for detachment of the ribosome. Such vectors may be obtained commercially or assembled from the sequences described by methods well known in the art.

Polynucleotides according to the invention may be  
20 inserted into the vectors described in an antisense orientation in order to provide for the production of antisense RNA. Antisense RNA or other antisense nucleic acids may be produced by synthetic means.

In accordance with the present invention, a  
25 defined nucleic acid includes not only the identical nucleic acid but also any minor base variations including in particular, substitutions in bases which result in a synonymous codon (a different codon specifying the same amino acid residue) due to the  
30 degenerate code. The term "nucleic acid sequence" also includes the complementary sequence to any single stranded sequence given regarding base variations.

The present invention also comprises within its scope proteins or polypeptides expressed by the  
35 nucleic acid molecules according to the invention or a functional equivalent, derivative or bioprecursor thereof.



- 7 -

The present invention also advantageously provides nucleic acid sequences of at least approximately 10 contiguous nucleotides of a nucleic acid according to the invention and preferably from 10 to approximately 120 nucleotides. In another aspect of the invention, nucleotide acid sequences are provided from 10 to 50 nucleotides. These sequences may, advantageously be used as probes or primers to initiate replication, or the like. Such nucleic acid sequences may be produced according to techniques well known in the art, such as by recombinant or synthetic means. They may also be used in diagnostic kits or the like for detecting the presence of a nucleic acid according to the invention. These tests generally comprise contacting the probe with the sample under hybridising conditions and detecting for the presence of any duplex or triplex formation between the probe and any nucleic acid in the sample.

According to the present invention, these probes may be anchored to a solid support. Preferably, they are present on an array so that multiple probes can simultaneously hybridize to a single biological sample. The probes can be spotted onto the array or synthesized *in situ* on the array. See Lockhart et al., Nature Biotechnology, Vol. 14, December 1996, "Expression monitoring by hybridization to high-density oligonucleotide arrays." A single array can contain more than up to more than a million different probes in discrete locations.

Advantageously, the nucleic acid sequences, according to the invention may be produced using such recombinant or synthetic means, such as for example using PCR cloning mechanisms which generally involve making a pair of primers, which may be between approximately 10 to 120 nucleotides to a region of the gene which is desired to be cloned, bringing the primers into contact with mRNA, cDNA, or genomic DNA

- 8 -

from a cell, performing a polymerase chain reaction under conditions which bring about amplification of the desired region, isolated the amplified region or fragment and recovering the amplified DNA. Generally, such techniques as defined herein are well known in the art, such as described in Sambrook et al (Molecular Cloning: a Laboratory Manual, 1989).

The nucleic acids or oligonucleotides according to the invention may carry a revealing label. Suitable labels include radioisotopes such as  $^{32}\text{P}$  or  $^{35}\text{S}$ , enzyme labels or other protein labels such as biotin or fluorescent markers. Such labels may be added to the nucleic acids or oligonucleotides of the invention and may be detected using known techniques *per se*.

The polypeptide or protein according to the invention includes all possible amino acid variants encoded by the nucleic acid molecule according to the invention including a polypeptide encoded by said molecule and having conservative amino acid changes. Polypeptides according to the invention further include variants of such sequences, including naturally occurring allelic variants which are substantially homologous to said polypeptides. In this context, substantial homology is regarded as a sequence which has at least 70%, preferably 80 or 90% amino acid homology with the polypeptides encoded by the nucleic acid molecules according to the invention.

Nucleic acids and polypeptides which are particularly preferred are those comprising the sequences of nucleotides illustrated in figures 1 to 3 and polypeptides illustrated in figures 14 to 16. However, a particularly preferred nucleic acid comprises the sequences of nucleotides illustrated in Figures 2 and/or 3, and their corresponding amino acid sequences identified in Figures 15 and 16.

Nucleotide sequences according to the invention

- 9 -

are particularly advantageous as selective therapeutic targets for treating *Candida albicans* associated infections. For example, an antisense nucleic acid capable of binding to the nucleic acid sequence  
5 illustrated in any of Figures 1 to 9 may be used to selectively inhibit expression of the corresponding polypeptides, leading to impaired growth of the *Candida albicans* with reductions of associated illnesses or diseases.

10 The nucleic acid molecule or the polypeptide according to the invention may be used as a medicament, or in the preparation of a medicament, for treating diseases or conditions associated with *Candida albicans* infection.

15 Advantageously, the nucleic acid molecule or the polypeptide according to the invention may be provided in a pharmaceutical composition together with a pharmaceutically acceptable carrier, diluent or excipient therefor.

20 The present invention is further directed to inhibiting expression of nucleic acids according to the invention *in vivo* by the use of antisense technology. Antisense technology can be used to control gene expression through triple-helix formation  
25 of antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion or the mature protein sequence, which encodes for the protein of the present invention, is used to design an antisense RNA  
30 oligonucleotide of from 10 to 50 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple-helix - see Lee et al. Nucl. Acids Res., 6:3073 (1979); Cooney et al., Science, 241:456 (1988);  
35 and Dervan et al., Science, 251: 1360 (1991), thereby preventing transcription and the production of the

- 10 -

corresponding protein. The antisense RNA oligonucleotide hybridises to the mRNA *in vivo* and blocks translation of an mRNA molecule into the corresponding protein (antisense - Okano, J.

5 Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)).

Antibodies to the protein or polypeptide of the present invention may, advantageously, be prepared by techniques which are known in the art. For example, 10 polyclonal antibodies may be prepared by inoculating a host animal, such as a mouse, with the polypeptide according to the invention or an epitope thereof and recovering immune serum. Monoclonal antibodies may be 15 prepared according to known techniques such as described by Kohler R. and Milstein C., Nature (1975) 256, 495-497.

Antibodies according to the invention may also be used in a method of detecting for the presence of a 20 polypeptide according to the invention, which method comprises reacting the antibody with a sample and identifying any protein bound to said antibody. A kit may also be provided for performing said method which comprises an antibody according to the invention and 25 means for reacting the antibody with said sample.

Proteins which interact with the polypeptide of the invention may be identified by investigating protein-protein interactions using the two-hybrid vector system first proposed by Chien et al. (1991).

30 This technique is based on functional reconstitution *in vivo* of a transcription factor which activates a reporter gene. More particularly the technique comprises providing an appropriate host cell with a DNA construct comprising a reporter gene under 35 the control of a promoter regulated by a transcription factor having a DNA binding domain and an activating domain, expressing in the host cell a first hybrid DNA

- 11 -

sequence encoding a first fusion of a fragment or all of a nucleic acid sequence according to the invention and either said DNA binding domain or said activating domain of the transcription factor, expressing in the host at least one second hybrid DNA sequence, such as a library or the like, encoding putative binding proteins to be investigated together with the DNA binding or activating domain of the transcription factor which is not incorporated in the first fusion; detecting any binding of the proteins to be investigated with a protein according to the invention by detecting for the presence of any reporter gene product in the host cell; optionally isolating second hybrid DNA sequences encoding the binding protein.

15       An example of such a technique utilises the GAL4 protein in yeast. GAL4 is a transcriptional activator of galactose metabolism in yeast and has a separate domain for binding to activators upstream of the galactose metabolising genes as well as a protein binding domain. Nucleotide vectors may be constructed, one of which comprises the nucleotide residues encoding the DNA binding domain of GAL4. These binding domain residues may be fused to a known protein encoding sequence, such as for example the nucleic acids according to the invention. The other vector comprises the residues encoding the protein binding domain of GAL4. These residues are fused to residues encoding a test protein. Any interaction between polypeptides encoded by the nucleic acid according to the invention and the protein to be tested leads to transcriptional activation of a reporter molecule in a GAL4 transcription deficient yeast cell into which the vectors have been transformed. Preferably, a reporter molecule such as  $\beta$ -galactosidase is activated upon restoration of transcription of the yeast galactose metabolism genes.

Further provided by the present invention is one

- 12 -

or more *Candida albicans* cells comprising an induced mutation in the DNA sequence encoding the polypeptide according to the invention.

5 A further aspect of the invention provides a method of identifying compounds which selectively inhibit or interfere with the expression, the functionality of polypeptides expressed from the nucleotides sequences illustrated in any of Figures 1 to 9 or the metabolic pathways in which these  
10 polypeptides are involved and which are critical for growth and survival of *Candida albicans*, which method comprises (a) contacting a compound to be tested with one or more *Candida albicans* cells having a mutation in a nucleic acid molecule according to the invention  
15 which mutation results in overexpression or underexpression of said polypeptides in addition to one or more wild type *Candida* cells, (b) monitoring the growth and/or activity of said mutated cell compared to said wild type wherein differential growth  
20 or activity of said one or more mutated *Candida* cells provides an indication of selective action of said compound on said polypeptide or another polypeptide in the same or a parallel pathway.

Compounds identifiable or identified using the  
25 method according to the invention, may advantageously be used as a medicament, or in the preparation of a medicament to treat diseases or conditions associated with *Candida albicans* infection. These compounds may also advantageously be included in a pharmaceutical  
30 composition together with a pharmaceutically acceptable carrier, diluent or excipient therefor.

A further aspect of the invention provides a method of identifying DNA sequences from a cell or organism which DNA encodes polypeptides which are  
35 critical for growth or survival, which method comprises (a) preparing a cDNA or genomic library from

- 13 -

said cell or organism in a suitable expression vector which vector is such that it can either integrate into the genome in said cell or that it permits transcription of antisense RNA from the nucleotide sequences in said cDNA or genomic library, (b) selecting transformants exhibiting impaired growth and determining the nucleotide sequence of the cDNA or genomic sequence from the library included in the vector from said transformant. Preferably, the cell or organism may be any yeast or filamentous fungus, such as, for example, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* or *Candida albicans*.

A further aspect of the invention provides a pharmaceutical composition comprising any of a compound, an antisense molecule or an antibody according to the invention together with a pharmaceutically acceptable carrier, diluent or excipient therefor.

The antisense molecules or indeed the compounds identified as agonists or antagonists of the nucleic acids or polypeptides according to the invention may be used in the form of a pharmaceutical composition, which may be prepared according to procedures well known in the art. Preferred compositions include a pharmaceutically acceptable vehicle or diluent or excipient, such as for example, a physiological saline solution. Other pharmaceutically acceptable carriers including other non-toxic salts, sterile water or the like may also be used. A suitable buffer may also be present allowing the compositions to be lyophilized and stored in sterile conditions prior to reconstitution by the addition of sterile water for subsequent administration. Incorporation of the polypeptides of the invention into a solid or semi-solid biologically compatible matrix may be carried out which can be implanted into tissues requiring treatment.

- 14 -

The carrier can also contain other pharmaceutically acceptable excipients for modifying other conditions such as pH, osmolarity, viscosity, sterility, lipophilicity, solubility or the like.

5 Pharmaceutically acceptable excipients which permit sustained or delayed release following administration may also be included.

The polypeptides, the nucleic acid molecules or compounds according to the invention may be  
10 administered orally. In this embodiment they may be encapsulated and combined with suitable carriers in solid dosage forms which would be well known to those skilled in the art.

As would be well known to those of skill in the  
15 art, the specific dosage regime may be calculated according to the body surface area of the patient or the volume of body space to be occupied, dependent upon the particular route of administration to be used. The amount of the composition actually  
20 administered will, however, be determined by a medical practitioner, based on the circumstances pertaining to the disorder to be treated, such as the severity of the symptoms, the composition to be administered, the age, weight, and response of the individual patient  
25 and the chosen route of administration.

The present invention may be more clearly understood with reference to the accompanying example, which is purely exemplary, with reference to the accompanying drawings, wherein

30

Figures 1 and 2: are nucleotide sequences isolated from *Candida albicans* and which have an identified function based on sequence homology with proteins from other organisms and which sequences are not present in the public domain.

35



- 15 -

Figures 3 : is a nucleotide sequence isolated from *Candida albicans* and which has an identified function based on sequence homology with proteins from other organisms and which sequence is partially present in the public domain.

10            Figures 4            :    is a nucleotide sequence of  
                                 previously unknown function  
                                 isolated from *Candida albicans* and  
                                 which is partially present in the  
                                 public domain.

15      Figures 5 to 9 : are nucleotide sequences of  
previously unknown function  
isolated from *Candida albicans*.

Figure 10 : is a diagrammatic representation  
of plasmid pGAL1PNiST-1.

Figure 11 : is a nucleotide sequence of plasmid pGAL1PNiST-1 of Figure 10.

25      Figure 12            :     is a diagrammatic representation  
                                     of plasmid pGALLPSiST-1.

Figure 13 : is a nucleotide sequence of plasmid pGAL1PSiST-1 of Figure 12.

Figures 14 to 20: are amino acid sequences of the appropriately corresponding DNA sequences illustrated in Figures 1 to 9 with reference to Table 1.

35      Figures 21 to 27: are growth curves of *Candida albicans* strains showing antisense

- 16 -

induced reduction in growth.

Figures 28 to 31: are growth curves of *Candida albicans* strains including knock-outs in the relevant gene identified.

### Example 1

Identification of novel drug targets in *C.*

10 *albicans* by anti-sense and disruptive integration

The principle of the approach is based on the fact that when a particular *C. albicans* mRNA is inhibited by producing the complementary anti-sense RNA, the corresponding protein will decrease. If this protein is critical for growth or survival, the cell producing the anti-sense RNA will grow more slowly or will die.

Since anti-sense inhibition occurs at mRNA level, the gene copy number is irrelevant, thus allowing applications of the strategy even in diploid organisms.

Anti-sense RNA is endogenously produced from an integrative or episomal plasmid with an inducible promoter; induction of the promoter leads to the production of an RNA encoded by the insert of the plasmid. This insert will differ from one plasmid to another in the library. The inserts will be derived from genomic DNA fragments or from cDNA to cover to the extent possible- the entire genome.

30 The vector is a proprietary vector allowing integration by homologous recombination at either the homologous insert or promoter sequence in the *Candida* genome. After introducing plasmids from cDNA or genomic libraries into *C. albicans*, transformants are screened for impaired growth after promoter (& thus anti-sense) induction in the presence of lithium

- 17 -

acetate. Lithium acetate prolongs the G1 phase and thus allows anti-sense to act during a prolonged period of time during the cell cycle. Transformants which show impaired growth in both induced and non-induced media, thus showing a growth defect due to integrative disruption, are selected as well.

Transformants showing impaired growth are supposed to contain plasmids which produce anti-sense RNA to mRNAs critical for growth or survival. Growth is monitored by measuring growth-curves over a period of time in a device (Bioscreen Analyzer, Labsystems) which allows simultaneous measurement of growth-curves of 200 transformants.

Subsequently plasmids can be recovered from the transformants and the sequence of their inserts determined, thus revealing which mRNA they inhibit. In order to be able to recover the genomic or cDNA insert which has integrated into the *Candida* genome, genomic DNA is isolated, cut with an enzyme which cuts only once into the library vector (and estimated approx. every 4096 bp in the genome) and religated. PCR with primers flanking the insert will yield (partial) genomic or cDNA inserts as PCR fragments which can directly be sequenced. This PCR analysis (on ligation reaction) will also show us how many integrations occurred. Alternatively the ligation reaction is transformed to *E. coli* and PCR analysis is performed on colonies or on plasmid DNA derived thereof.

This method is employed for a genome-wide search for novel *C. albicans* genes which are important for growth or survival.

#### Materials & Methods

##### Constructi n of pGallPNiST-1

The backbone of the pGAL1PNiST-1 vector (integrative anti-sense *Sfi*I-*Not*I vector) is

- 18 -

pGEM11zf(+) (Promega Inc.). First, the CaMAL2 EcoRI/SalI promoter fragment from pDBV50 (D.H. Brown et al. 1996) was ligated into EcoRI/SalI-opened pGEM11zf(+) resulting in the intermediate construct pGEMMAL2P-1. Into the latter (MscI/CIP) the CaURA3 selection marker was cloned as a Eco47III/XmnI fragment derived from pRM2. The resulting pGEMMAL2P-2 vector was NotI/HindIII opened in order to accept the NotI-stuffer-SfiI cassette from pPCK1NiSCYCT-1 (EagI/HindIII fragment): pMAL2PNiST-1. Finally, the plasmid pGAL1PNiST-1 was constructed by exchanging the SalI/Ecl136II MAL2 promoter in pMAL2PNiST-1 by the XhoI/SmaI GAL1 promoter fragment derived from pRM2GAL1P.

15

#### Construction of pGAL1PSiST-1

The vector pGAL1PSiST-1 was created for cloning the small genomic DNA fragments (flanked by SfiI sites) behind the GAL1 promoter. The only difference with pGAL1PNiST-1 is that the hIFN $\beta$  (stuffer fragment) insert fragment in pGAL1PSiST-1 is flanked by two SfiI sites instead of a SfiI and a NotI site as in pGAL1PNiST-1. To construct pGAL1PSiST-1 the EcoRI-HindIII fragment, containing hIFN $\beta$  flanked by a SfiI and a NotI site, of pMAL2pHiET-3 (unpublished) was exchanged by the EcoRI-HindIII fragment, containing hIFN $\beta$  flanked by two SfiI sites, from YCp50S-S (an *E. coli* / *S. cerevisiae* shuttle vector derived from the plasmid YCp50, which is deposited in the ATCC collection (number 37419; Thrash et al., 1985); an EcoRI-HindIII fragment, containing the gene hIFN $\beta$ , which is flanked by two SfiI sites, was inserted in YCp50, creating YCp50S-S), resulting into plasmid pMAL2PSiST-1. The MAL2 promoter from pMAL2PSiST-1 (by a NaeI-balI digest) was further replaced by the GAL1 promoter from pGAL1PNiST-1 (via a XhoI-FSPI digest),

35

- 19 -

creating the vector pGAL1PSiST-1.

***Candida albicans* genomic library**

**\* Preparation of the genomic DNA fragments**

5 A *Candida albicans* genomic DNA library with small DNA fragments (400 to 1,000 bp) was prepared. Genomic DNA of *Candida albicans* B2630 was isolated following a modified protocol of Blin and Stafford (1976). The quality of the isolated genomic DNA was checked by gel  
10 electrophoresis. Undigested DNA was located on the gel above the marker band of 26,282 bp. A little smear, caused by fragmentation of the DNA, was present. To obtain enrichment for genomic DNA fragments of the desired size, the genomic DNA was partially digested.  
15 Several restriction enzymes (*AluI*, *HaeIII* and *RsaI*; all creating blunt ends) were tried out. The appropriate digest conditions have been determined by titration of the enzyme. Enrichment of small DNA fragments was obtained with 70 units of *AluI* on 10 µg  
20 of genomic DNA for 20 min. T4 DNA polymerase (Boehringer) and dNTPs (Boehringer) were added to polish the DNA ends. After extraction with phenol-chloroform the digest was size-fractionated on an agarose gel. The genomic DNA fragments with a length  
25 of 500 to 1,250 bp were eluted from the gel by centrifugal filtration (Zhu et al., 1985). *SfiI* adaptors (5' GTTGGCCTTTT) or (5' AGGCCAAC) were attached to the DNA ends (blunt) to facilitate cloning of the fragments into the vector. Therefore, a 8-mer  
30 and 11-mer oligonucleotide (comprising the *SfiI* site) were kinased and annealed. After ligation of these adaptors to the DNA fragments a second size-fractionation was performed on an agarose gel. The DNA fragments of 400 to 1150 bp were eluted from the  
35 gel by centrifugal filtration.

**\* Preparation of the pGAL1PSiST-1 vector fragment**

- 20 -

The small genomic DNA fragments were cloned after the GAL1 promoter in the vector pGAL1PSiST-1. Qiagen-purified pGAL1PSiST-1 plasmid DNA was digested with SfiI and the largest vector fragment eluted from the gel by centrifugal filtration (Zhu et al., 1985).  
5 Ligation with a control DNA fragment, flanked by SfiI sites, was performed as a control. The ligation mix was electroporated to MC1061 *E. coli* cells. Plasmid DNA of 24 clones was analyzed. In all cases the  
10 control fragment was inserted in the pGAL1PSiST-1 vector fragment.

*\* Upscaling*

All genomic DNA fragments (450 ng) were ligated into the pGAL1PSiST-1 vector (20 ng). After  
15 electroporation at 2500V, 40pF circa 400,000 clones were obtained. These clones were pooled into three groups and stored as glycerol slants. Also Qiagen-purified DNA was prepared from these clones. A clone  
20 analysis showed an average insert length of 600 bp and a percentage of 91 for clones with an insert. The size of the library corresponds to 5 times the diploid genome. The genomic DNA inserts are sense or anti-sense orientated in the vector.

25 ***Candida albicans* cDNA library**

Total RNA was extracted from *Candida albicans* B2630 grown on respectively minimal (SD) and rich (YPD) medium as described by Chirgwin et al. in  
Sambrook et al 1996. mRNA was prepared from total RNA  
30 using the Invitrogen Fast Track procedure.

First strand cDNA is synthesised with the Superscript Reverse Transcriptase (BRL) and with an oligo dT-NotI Primer adapter. After second strand synthesis, cDNA is polished with Klenow enzyme and  
35 purified over a Sephacryl S-400 spun column. Phosphorylated SfiI adapters are then ligated to the

- 21 -

cDNA, followed by digestion with the NotI restriction enzyme. The SfiI/NotI cDNA is then purified and sized on a Biogel column A150M.

5 First fraction contains approximately 38,720 clones by transformation, the second fraction only 1540 clones. Clone analysis:  
Fr. I: 22/24 inserts, 16 ' 1000 bp, 4 ' 2000 bp, average size: 1500 bp.  
Fr. II: 9/12 inserts, 3 ' 1000 bp, average size: 960  
10 bp cDNA was ligated in a NotI/SfiI opened pGAL1PNiST-1 vector (anti-sense)

#### **Candida transformation**

15 The host strain used for transformation is a *C. albicans* ura3 mutant, CAI-4, which contains a deletion in orotidine-5'-phosphate decarboxylase and was obtained from William Fonzi, Georgetown University (Fonzi and Irwin). CAI-4 was transformed with the above described cDNA library or genomic library using  
20 the *Pichia* spheroplast module (Invitrogen). Resulting transformants were plated on minimal medium supplemented with glucose (SD, 0.67% or 1.34% Yeast Nitrogen base w/o amino acids + 2% glucose) plates and incubated for 2-3 days at 30°C.

25

#### **Screening for mutants**

Starter cultures were set up by inoculating each colony in 1 ml SD medium and incubating overnight at 30°C and 300 rpm. Cell densities were determined using  
30 a Coulter counter (Coulter Z1; Coulter electronics limited). 250.000 cells/ml were inoculated in 1 ml SD medium and cultures were incubated for 24 hours at 30°C and 300 rpm. Cultures were washed in minimal medium without glucose (S) and the pellet resuspended  
35 in 650 µl S medium. 8 µl of this culture is used for inoculating 400 µl cultures in a Honeywell-100 plate

- 22 -

(Bioscreen analyzer; Labsystems). Each transformant was grown during three days in S medium containing LiAc; pH 6.0, with 2% glucose/2% maltose or 2% galactose/2% maltose respectively while shaking every 3 minutes for 20 seconds. Optical densities were measured every hour during three consecutive days and growth curves were generated (Bioscreen analyzer; Labsystems).

Growth curves of transformants grown in respectively anti-sense non-inducing (glucose/maltose) and inducing (galactose/maltose) medium are compared and those transformants showing impaired growth upon anti-sense induction are selected for further analysis. Transformants showing impaired growth by virtue of integration into a critical gene are also selected.

#### Isolation of genomic or cDNA inserts

Putatively interesting transformants are grown in 1.5 ml SD overnight and genomic DNA is isolated using the Nucleon MI Yeast kit (Clontech). Concentration of genomic DNA is estimated by analyzing a sample on an agarose gel.

20 ng of genomic DNA is digested for three hours with an enzyme that cuts uniquely in the library vector (SacI for the genomic library; PstI for the cDNA library) and treated with RNase. Samples are phenol/chloroform extracted and precipitated using NaOAc/ethanol.

The resulting pellet is resuspended in 500  $\mu$ l ligation mixture (1 x ligation buffer and 4 units of T4 DNA ligase; both from Boehringer) and incubated overnight at 16°C.

After denaturation (20 min 65°C), purification (phenol/chloroform extraction) and precipitation (NaOAc/ethanol) the pellet is resuspended in 10  $\mu$ l MilliQ (Millipore) water.



- 23 -

**PCR analysis**

Inverse PCR is performed on 1  $\mu$ l of the precipitated ligation reaction using library vector specific primers (oligo23 5' TGC-AGC-TCG-ACC-TCG-ACT-G 3' and oligo25 5' GCG-TGA-ATG-TAA-GCG-TGA-C 3' for the genomic library; 3pGALNistPCR primer :5'TGAGCAGCTCGCCGTCGCGC 3' and 5pGALNistPCR primer: 5'GAGTTATACCCTGCAGCTCGAC 3' for the cDNA library; both from Eurogentec) for 30 cycles each consisting of (a) 1 min at 95 °C, (b) 1 min at 57 °C, and (c) 3 min at 72 °C. In the reaction mixture 2.5 units of Taq polymerase (Boehringer) with TaqStart antibody (Clontech) (1:1) were used, and the final concentrations were 0.2  $\mu$ M of each primer, 3 mM MgCl<sub>2</sub> (Perkin Elmer Cetus) and 200  $\mu$ M dNTPs (Perkin Elmer Cetus). PCR was performed in a Robocycler (Stratagene).

**Sequence determination**

Resulting PCR products were purified using PCR purification kit (Qiagen) and were quantified by comparison of band intensity on EtBr stained agarose gel with the intensity of DNA marker bands. The amount of PCR product (expressed in ng) used in the sequencing reaction is calculated as the length of the PCR product in basepairs divided by 10. Sequencing reactions were performed using the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit according to the instructions of the manufacturer (PE Applied Biosystems, Foster City, CA) except for the following modifications.

The total reaction volume was reduced to 15  $\mu$ l. Reaction volume of individual reagents were changed accordingly. 6.0  $\mu$ l Terminator Ready Reaction Mix was replaced by a mixture of 3.0  $\mu$ l Terminator Ready Reaction Mix + 3.0  $\mu$ l Half Term (GENPAK Limited,

- 24 -

Brighton, UK). After cycle sequencing, reaction mixtures were purified over Sephadex G50 columns prepared on Multiscreen HV opaque microtiter plates (Millipore, Molsheim, Fr) and were dried in a speedVac. Reaction products were resuspended in 3  $\mu$ l loading buffer. Following denaturation for 2 min at 95°C, 1  $\mu$ l of sample was applied on a 5% Long Ranger Gel (36 cm well-to-read) prepared from Singel Packs according to the supplier's instructions (FMC BioProducts, Rockland, ME). Samples were run for 7 hours 2X run on a ABI 377XL DNA sequencer. Data collection version 2.0 and Sequence analysis version 3.0 (for basecalling) software packages are from PE Applied Biosystems. Resulting sequence text files were copied onto a server for further analysis.

#### Sequence analysis

Nucleotide sequences were imported in the VectorNTI software package (InforMax Inc, North Bethesda, MD, USA), and the vector and insert regions of the sequences were identified. Sequence similarity searches against public and commercial sequence databases were performed with the BLAST software package (Altschul et al., 1990) version 1.4. Both the original nucleotide sequence and the six-frame conceptual translations of the insert region were used as query sequences. The used public databases were the EMBL nucleotide sequence database (Stoesser et al., 1998), the SWISS-PROT protein sequence database and its supplement TrEMBL (Bairoch and Apweiler, 1998), and the ALCES *Candida albicans* sequence database (Stanford University, University of Minnesota). The commercial sequence databases used were the LifeSeq® human and PathoSeq® microbial genomic databases (Incyte Pharmaceuticals Inc., Palo Alto, CA, USA), and the GENESEQ patent sequence database (Derwent, London,

- 25 -

UK). Three major results were obtained on the basis of the sequence similarity searches: function, novelty, and specificity. A putative function was deduced on the basis of the similarity with sequences with a known function, the novelty was based on the absence or presence of the sequences in public databases, and the specificity was based on the similarity with vertebrate homologues.

## 10           **Methods**

Blastx of the nucleic acid sequences against the appropriate protein databases: Swiss-Prot for clones of which the complete sequence is present in the public domain, and paorfp (PathoSeq™) for clones of which the complete sequences is not present in the public domain.

The protein to which the translated nucleic acid sequence corresponds to is used as a starting point. The differences between this protein and our translated nucleic acid sequences are marked with a double line and annotated above the protein sequence. The following symbols are used:

a one-letter amino acid code or the ambiguity code X is used if our translated nucleic acid sequence has another amino acid on a certain position,

the stop codon sign \* is used if our translated nucleic acid sequence has a stop codon on a certain position,

The letters fs (frame shift) are used if a frame shift occurs in our translated nucleic acid sequence, and another reading frame is used,

the words ambiguity or ambiguities are used if a part of our translated nucleic acid sequence is present in the proteins, but not visible in the alignments of the blast results,

The phrase "missing sequence" is used if the translated nucleic acid sequence does not comprise

- 26 -

that part of the protein.

Blastx: compares the six-frame conceptual translation products of a nucleotide query sequence (both strands) against a protein sequence database.

5

#### Gene Knock-outs

To verify that the growth effect was due to the interference with the identified gene and to support the specificity of the antisense effect, single allele knock-outs were made in the identified genes (Figures 28 to 31) using the URA-blaster method (Fonzi and Irwin 1993).

#### 15 Screening for compounds modulating expression of polypeptides critical for growth and survival of *C. albicans*

The method proposed is based on observations (Sandbaken et al., 1990; Hinnebusch and Liebman 1991; Ribogene PCT WO 95/11969, 1995) suggesting that underexpression or overexpression of any component of a process (e.g. translation) could lead to altered sensitivity to an inhibitor of a relevant step in that process. Such an inhibitor should be more potent against a cell limited by a deficiency in the macromolecule catalyzing that step and/or less potent against a cell containing an excess of that macromolecule, as compared to the wild type (WT) cell.

Mutant yeast strains, for example, have shown that some steps of translation are sensitive to the stoichiometry of macromolecules involved. (Sandbaken et al. 1996). Such strains are more sensitive to compounds which specifically perturb translation (by acting on a component that participates in translation) but are equally sensitive to compounds with other mechanisms of action.

35

This method thus not only provides a means to

- 27 -

identify whether a test compound perturbs a certain process but also an indication of the site at which it exerts its effect. The component which is present in altered form or amount in a cell whose growth is  
5 affected by a test compound is potentially the site of action of the test compound.

The assay to be set up involves measurement of growth of an isogenic strain which has been modified only in a certain specific allele, relative to a wild  
10 type (WT) *C. albicans* strain, in the presence of R-compounds. Strains can be ones in which the expression of a specific essential protein is impaired upon induction of anti-sense or strains which carry  
15 disruptions in an essential gene. An *in silico* approach to finding novel essential genes in *C. albicans* will be performed. A number of essential genes identified in this way will be disrupted (in one allele) and the resulting strains can be used for  
20 comparative growth screening.

#### Assay for High Throughput screening for drugs

35  $\mu$ l minimal medium (S medium + 2% galactose + 2% maltose) is transferred in a transparent flat-bottomed 96 well plate using an automated pipetting  
25 system (Multidrop, Labsystems). A 96-channel pipettor (Hydra, Robbins Scientific) transfers 2.5  $\mu$ l of R-compound at  $10^{-3}$  M in DMSO from a stock plate into the assay plate.

The selected *C. albicans* strains (mutant and  
30 parent (CAI-4) strain) are stored as glycerol stocks (15%) at  $-70^{\circ}\text{C}$ . The strains are streaked out on selective plates (SD medium) and incubated for two days at  $30^{\circ}\text{C}$ . For the parent strain, CAI-4, the medium is always supplemented with 20  $\mu\text{g/ml}$  uridine. A single  
35 colony is scooped up and resuspended in 1 ml minimal medium (S medium + 2% galactose + 2% maltose). Cells

- 28 -

are incubated at 30°C for 8 hours while shaking at 250 rpm. A 10 ml culture is inoculated at 250.000 cells/ml. Cultures are incubated at 30°C for 24 hours while shaking at 250 rpm. Cells are counted in  
5 Coulter counter and the final culture (S medium + 2% galactose + 2% maltose) is inoculated at 20.000 to 50.000 cells/ml. Cultures are grown at 30°C while shaking at 250 rpm until a final OD of 0.24 (+/- 0.04) 600nm is reached.

10 200 µl of this yeast suspension is added to all wells of MW96 plates containing R-compounds in a 450 (or 250) µl total volume. MW96 plates are incubated (static) at 30°C for 48 hours.

Optical densities are measured after 48 hours.  
15 Test growth is expressed as a percentage of positive control growth for both mutant (x) and wild type (y) strains. The ratio (x/y) of these derived variables is calculated.

- 29 -

References

- 5 Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and  
D.J. Lipman, 1990, Basic Local Alignment Search Tool,  
J. Mol. Biol. 215: 403-410.
- Thrash C., Bankier A. T., Barrell B. G. and Sternglanz  
R. (1985) *Proc. Natl. Acad. Sci. USA* 82: 4374-4378.
- 10 Blin N. and Stafford D. W. (1976) *Nucleic Acids Res.*  
3: 2303-2308.
- Zhu J., Kempnaers W., Van Der Straeten D., Contreras  
R., and Fiers W. (1985) *Bio/Technology* 3: 1014-1016.
- 15 Sambrook J., Fritsch E.F. and Maniatis T. (1989)  
*Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold  
Spring Harbor Laboratory, Cold Spring Harbor, NY.
- 20 Fonzi W. and Irwin H. (1993) *Genetics* 134:717-728.  
Altschul S.F, Gish W., Miller W., Myers E.W., Lipman  
D.J. (1990) *J. Mol. Biol.* 215(3):403-410.
- Bairoch A. and Apweiler R. (1998) *Nucleic Acids Res.*  
25 26(1):38-42.
- Stoesser G., Moseley M.A., Sleep J., McGowran M.,  
Garcia-Pastor M., Sterk P. (1998) *Nucleic Acids Res.*  
26(1):8-15.
- 30 Chien et al., (1991) *Proc. Natl. Acad. Sci USA* 88,  
9578-9582.
- Sandbaken M.G. Lupisella J.A., DiDomenico B. and  
35 Chakraborty K., *J. Biol. Chem.* 265:15838-15844, 1990.

- 30 -

Hinnebusch A.G. and Liebman S.W., in : The Molecular Biology of the Yeast *Saccharomyces*, Broach J.R., Pringle J.R. and Jones E.W., eds., CSH Laboratory Press; NY 1991.

5

Patent application RiboGene Inc., PCT WO 95./11969, 1995.

10

Brown, D.H. et al., Molecular General Genetics 251 (1) 75-80, 1996.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215: 403-410.

15

Arndt, G.M., D. Atkins, M. Patrikakis, and J.G. Izant 1995. Gene regulation by antisense RNA in the fission yeast *Schizosaccharomyces pombe*. Mol. Gen. Genet. 248: 293-300.

20

Atkins, D., and W.L. Gerlach. 1994. Artificial ribozyme and antisense gene expression in *S. cerevisiae*. Antisense research and development 4:109-117.

25

Atkins, D., G.M. Arndt, and J.G. Izant 1994. Antisense gene expression in yeast. Biol. Chem. Hoppe-Seyler. 375:721-729.

30

Atkins, D., M. Patrikakis, J.G. Izant 1995. The *ade6* gene of the fission yeast as a target for antisense and ribozyme RNA-mediated suppression. Antisense Research & Development. 5(4):295-305.

35

Bairoch, A., and R. Apweiler. 1998. The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1998. Nucleic Acids Res. 26: 38-42.



- 31 -

- Baudin, A., O. Ozier-Kalogeropoulos, A. Denouel, F. Lacroute, C. Cullin. 1993. A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Research*. 21(14):3329-30.
- 5
- Blin, N., and D.W. Stafford. 1976. *Nucleic Acids Res.* 3: 2303-2308.
- Dujon, B. 1998. European Functional Analysis Network (EUROFAN) and the functional analysis of the *Saccharomyces cerevisiae* genome. *Electrophoresis*. 19:617-624.
- 10
- Del Rosario, M., J.C. Stephans, J. Zakel, J. Escobedo, and K. Giese. 1996. Positive selection system to screen for inhibitors of human immunodeficiency virus-1 transcription. *Nature Biotechnology*. 14:1592-1596.
- 15
- Fairhead, C., A. Thierry, F. Denis, M. Eck, and B. Dujon. 1998. "Mass-murder" of ORFs from three regions of chromosome XI from *Saccharomyces cerevisiae*. *Gene*. 223:33-46.
- 20
- Ferbeyre, G., J. Bratty, H. Chen, R. Cedergren. 1996. Cell cycle arrest promotes trans-hammerhead ribozyme action in yeast. *Journal of Biological Chemistry*. 271(32):19318-23.
- 25
- Fleischmann, R.D., M.D. Adams, O. White, R.A. Clayton, E.F. Kirkness, A.R. Kerlavage, C.J. Bult, J.F. Tomb, B.A. Dougherty, J.M. Merrick. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae*. *Science*. 269: 496-512
- 30
- Fonzi, W.A., and M.Y. Irwin. 1993. Isogenic strain construction and gene mapping in *Candida albicans*. *Genetics* 134:717-728.
- 35

- 32 -

Fraser, C.M., J.D. Gocayne, O. White, M.D. Adams, R.A. Clayton, R.D. Fleischmann, C.J. Bult, A.R. Kervalage, G. Sutton, J.M. Kelley. 1995. The minimal gene complement of *Mycoplasma genitalium*. Science 270:397-403.

Giese, K. 1997. Method and construct for screening for inhibitors of transcriptional activation. International patent application WO 97/10360.

Hahn, S. Guarente L. 1988. Yeast HAP2 and HAP3: transcriptional activators in a heteromeric complex. Science. 240(4850):317-21

Heid, C.A., J. Stevens, K.J. Livak, and P.M. Williams. 1996. Real time quantitative PCR. Genome Methods 6:986-994.

Jayaram, M., A. Sutton, J.R. Broach. 1985. Properties of REP3: a cis-acting locus required for stable propagation of the *Saccharomyces cerevisiae* plasmid 2 microns circle. Molecular & Cellular Biology. 5(9):2466-75.

Jeong, S.W., W.H. Lang, R.H. Reeder. 1996. The yeast transcription terminator for RNA polymerase I is designed to prevent polymerase slippage. Journal of Biological Chemistry. 271(27):16104-10.

Johnson, F.B., D.A. Sinclair, and L. Guarente. 1999. Molecular Biology of aging. Cell. 96:291-302.

Leuker, C.E., A. Sonneborn, S. Delbruck, J.F. Ernst. 1997. Sequence and promoter regulation of the PCK1 gene encoding phosphoenolpyruvate carboxykinase of the fungal pathogen *Candida albicans*. Gene. 192(2):235-40.

- 33 -

- Lie, Y.S., and C.J. Petropoulos. 1998. Advances in quantitative PCR technology: 5' nuclease assays. *Current Opinion in Biotechnology* 9:43-48.
- 5 Livak, K.J., S.J. Flood, J. Marmaro, W. Giusti, K. Deetz. 1995. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *Genome Research*. 4(6):357-62.
- 10 Mandart, E. 1998. Effects of mutations in the *Saccharomyces cerevisiae* RNA14 gene on the abundance and polyadenylation of its transcripts. *Mol. Gen. Genet.* 258:16-25.
- 15 McMahon, S.B., H.A. Van Buskirk, K.A. Dugan, T.D. Copeland, and M.D. Cole. 1998. The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. *Cell*. 94:363-74.
- 20 Murray, J.A.H., M. Scarpa, N. Rossi, G. Cesareni. 1987. Antagonistic controls regulate copy number of the yeast 2 $\mu$  plasmid. *EMBO J.* 6:4205-4212.
- 25 Nasr, F., A. Bécam, P.P. Slonimski, and C.J. Herbert. 1994. YBR1012, an essential gene from *S. cerevisiae*: construction of an RNA-antisense conditional allele and isolation of a multicopy suppressor. *CR Acad. Sci. Paris*. 317:607-613
- 30 Nasr, F., A. Bécam, S.C. Brown, D. De Nay. P.P. Slonimski, and C.J. Herbert. 1995. Artificial antisense regulation of YBR1012 an essential gene from *S. cerevisiae* which is important for progression through G1/S. *Mol. Gen. Genet.* 249:51-57.
- 35

- 34 -

- Nomura, T., N. Fujita, A. Ishihama. 1985. Promoter selectivity of *E. coli* RNA polymerase: analysis of the promoter system of convergently-transcribed *dnaQ-rnh* genes. *Nucleic Acids Research*. 13(21):7647-61.
- 5 Orlando, C., P. Pinzani, and M. Pazzagli. 1998. Developments in quantitative PCR. *Clin. Chem. Lab. Med.* 36(5):255-269.
- 10 Pla, J., C. Gil, F. Monteoliva, M. Sanchez, and C. Nombela. 1996. Understanding *Candida albicans* at the molecular level. *Yeast*. 12:1677-1702.
- 15 Reeder, R.H. and W.H. Lang. 1997. Terminating transcription in eukaryotes: lessons learned from RNA polymerase I. *Trends in Biochemical Sciences*. 22(12):473-7, 1997
- 20 Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Sinclair, D.A., L. Guarente. 1997. Extrachromosomal rDNA circles--a cause of aging in yeast. *Cell*. 25 91(7):1033-42.
- 30 Smith, V., D. Botstein, and P. O. Brown. 1995. Genetic footprinting: A genomic strategy for determining a gene's function given its sequence. *Proc. Natl. Acad. Sci. USA*. 92:6479-6483.
- 35 Stoesser, G., Moseley M.A., Sleep J., McGowran M., Garcia-Pastor M., Sterk P. 1998. *Nucleic Acids Res.* 26(1):8-15.
- Thompson-Jager, S. Domdey H. 1990. The intron of the yeast actin gene contains the promoter for an

- 35 -

antisense RNA. *Current Genetics*. 17(3):269-73.

Thompson, J.R., E. Register, J. Curotto, M. Kurtz, and  
R. Kelly. 1998. An improved protocol for the  
5 preparation of yeast cells for transformation by  
electroporation. *Yeast*. 14:565-571.

Thrash, C., A.T. Bankier, B.G. Barrell, and R.  
Sternglanz. 1985. *Proc. Natl. Acad. Sci. USA* 82: 4374-  
10 4378.

Van Duin, M., J. van Den Tol, J.H. Hoeijmakers, D.  
Bootsma, I.P. Rupp, P. Reynolds, L. Prakash, and S.  
Prakash. 1989. Conserved pattern of antisense  
15 overlapping transcription in the homologous human  
ERCC-1 and yeast RAD10 DNA repair gene regions.  
*Molecular & Cellular Biology*. 9(4):1794-8.

Wach, A., A. Brachet, R. Pohlmann, P. Philippsen.  
20 1994. New heterologous modules for classical or PCR-  
based gene disruptions in *Saccharomyces cerevisiae*.  
*Yeast*. 10(13):1793-808.

Wilson, R.B., D. Davis, A.P. Mitchell. 1999. Rapid  
25 hypothesis testing with *Candida albicans* through gene  
disruption with short homology regions. *Journal of*  
*Bacteriology*. 181(6):1868-74.

Zhu, J., W. Kempenaers, D. Van der Straeten, R.  
30 Contreras, and W. Fiers. 1985. A method for fast and  
pure DNA elution from agarose gels by centrifugal  
filtration. *Bio/Technology*. 3: 1014-1016.

- 36 -

**Table 1**

Seq ID No.	Figure No.	Clone	Function
1	1	382c_cp	-
2	2	392c_cp	TUF1
3	3	-	RAD53
4	4	417c_cpG2L	-
5	5	323c_af	-
6	6	322c_cp <sup>1</sup>	-
7	7	26g3	-
8	8	409c_cp	-
9	9	382c_cpG1L2	-
10	14	382c_cp (prt)	-
11	15	392c_cp (prt)	TUF1
12	16		RAD53
13	17	325c_af (prt) <sup>2</sup>	-
14	18	322c_cp (prt) <sup>2</sup>	-
15	19	26g3 (prt)	-
16	20	417c_cp 92L (prt)	-

- 322c-cp is a member of the UPF0057 protein family. It contains potential transmembrane regions (6-23aa; 30-53aa) and could be low temperature or salt-stress inducible.
- 325c-af shows similarity to IMP4 yeast and related proteins and it might be involved in rRNA processing in *Candida albicans* in a similar way to IMP4.

- 37 -

Claims

1. A nucleic acid molecule encoding a polypeptide which is critical for survival and growth of the yeast *Candida albicans* and which nucleic acid molecule comprises any of the sequences of nucleotides illustrated in Seq ID Nos 1 to 9.
2. A nucleic acid molecule encoding a polypeptide which is critical for survival and growth of the yeast *Candida albicans* and which nucleic acid molecule comprises any of the sequences of nucleotides illustrated in Seq ID Nos 1 to 3.
3. A nucleic acid molecule encoding a polypeptide which is critical for survival and growth of the yeast *Candida albicans* and which nucleic acid molecule comprises any of the sequences of nucleotides illustrated in Seq ID Nos 1 or 2 and fragments or derivatives of said nucleic acid molecules.
4. A nucleic acid molecule according to any of claims 1 to 3 which is mRNA.
5. A nucleic acid molecule according to any of claims 1 to 3 which is DNA.
6. A nucleic acid molecule according to claim 5 which is cDNA.
7. A nucleic acid molecule capable of hybridising to the molecules according to any of claims 1 to 6 or the sequences illustrated in any of Seq ID Nos 1 to 9 under high stringency conditions.
8. An antisense molecule comprising a nucleic

- 38 -

acid molecule capable of hybridising to the molecules according to any of claims 1 to 6 or the sequences illustrated in any of Seq ID Nos 1 to 9.

5           9. Cells containing a nucleic acid molecule according to any of claims 1 to 8, wherein said cells are bacterial or eukaryotic.

10           10. A polypeptide encoded by the nucleic acid molecule according to any of claims 1 to 7 or the sequences illustrated in any of Seq ID Nos 1 to 9.

15           11. A polypeptide having any of amino acid sequences illustrated in any of Seq ID Nos 14 to 20.

            12. A recombinant DNA construct comprising a nucleic acid molecule according to claim 5 or 6.

20           13. A recombinant DNA construct comprising a nucleic acid molecule according to claim 5 or 6 wherein said nucleic acid molecule is inserted in the antisense orientation.

25           14. A recombinant DNA construct according to claim 12 or 13 wherein said recombinant DNA construct is an expression vector.

30           15. A construct according to claim 14 which comprises an inducible promoter.

            16. A construct according to claim 14 or 15 which comprises a sequence encoding a reporter molecule.

35           17. Cells containing a recombinant DNA construct according to any of claims 12 to 16, wherein said cells are bacterial or eukaryotic.



- 39 -

18. A nucleic acid molecule according to any of claims 1 to 8 or the nucleotide sequences illustrated in Seq ID Nos 1 to 9 for use as a medicament.

5           19. Use of a nucleic acid molecule according to any of claims 1 to 8 or the sequences illustrated in Seq ID Nos 1 to 9 in the preparation of a medicament for treating *Candida albicans* associated diseases.

10           20. A polypeptide according to claim 10 or 11 for use as a medicament.

            21. Use of a polypeptide according to claim 10 or 11 in the preparation of a medicament for treating  
15   *Candida albicans* associated infections.

            22. A pharmaceutical composition comprising a nucleic acid molecule according to any of claims 1 to 8 or a polypeptide according to claim 10 or 11  
20   together with a pharmaceutically acceptable carrier diluent or excipient therefor.

            23. A *Candida albicans* cell comprising an induced mutation in the DNA sequence encoding the  
25   polypeptide according to claim 10.

            24. A method of identifying compounds which selectively modulate expression or functionality of polypeptides or metabolic pathways in which these  
30   polypeptides are involved and which are crucial for growth and survival of *Candida albicans*, which method comprises:

            (a) contacting a compound to be tested with one or more *Candida albicans* cells having a  
35   mutation in a nucleic acid molecule according to any of claims 1 to 8 which

- 40 -

mutation results in overexpression or underexpression of said polypeptides in addition to contacting one or more wild type *Candida albicans* cells with said compound,

5 (b) monitoring the growth and/or activity of said mutated cell compared to said wild type; wherein differential growth or activity of said one or more mutated *Candida* cells is indicative of selective action of

10 said compound on a polypeptide or another polypeptide in the same or a parallel pathway.

25. A compound identifiable according to the

15 method of claim 24.

26. A compound according to claim 25 for use as a medicament.

20 27. Use of a compound according to claim 25 in the preparation of a medicament for treating *Candida albicans* associated diseases.

28. A pharmaceutical composition comprising a

25 compound according to claim 25 together with a pharmaceutically acceptable carrier, diluent or excipient therefor.

29. A method of identifying DNA sequences from a

30 cell or organism which DNA encodes polypeptides which are critical for growth or survival of said cell or organism, which method comprises:

(a) preparing a cDNA or genomic library from said cell or organism in a suitable

35 expression vector which vector is such that it can either integrate into the genome in said cell or that it permits transcription

- 41 -

of antisense RNA from the nucleotide sequences in said cDNA or genomic library.

- (b) selecting transformants exhibiting impaired growth and determining the nucleotide sequence of the cDNA or genomic sequence from the library included in the vector from said transformant.

5

30. A method according to claim 29 wherein said cell or organism is a yeast or filamentous fungus.

10

31. A method according to claim 29 or 30 wherein said cell or organism is any of *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* or *Candida albicans*.

15

32. An antibody capable of binding to a polypeptide according to claim 10 or 11.

20

33. An oligonucleotide comprising a fragment of from 10 to 120 contiguous nucleotides of a nucleic acid molecule according to any of claims 1 to 8.

25

34. An oligonucleotide according to claim 33 comprising a fragment of from 10 to 50 contiguous nucleotides.

30

35